

## **Methodological Framework for Evaluating Intelligent Transportation Systems: COMPASS4D**

**Simon Edwards<sup>1</sup>, Graeme Hill<sup>1</sup>, Evangelos Mitsakis<sup>2</sup>, Josep Maria Salanova Grau<sup>3</sup>, Jaap Vreeswijk<sup>4</sup>, Jose Manuel Martínez<sup>5</sup>, Guillaume Vernet<sup>6</sup>**

<sup>1</sup>Transport Operations Research group, University of Newcastle upon Tyne, Phone: +44 191 222 8117, E-mail: [simon.edwards@ncl.ac.uk](mailto:simon.edwards@ncl.ac.uk); [g.a.hill@ncl.ac.uk](mailto:g.a.hill@ncl.ac.uk)

<sup>2</sup>Centre for Research and Technology Hellas/Hellenic Institute of Transport, Thessaloniki, Greece, Phone: +30 2310 498 459, E-mail: [emit@certh.gr](mailto:emit@certh.gr)

<sup>3</sup>Centre for Research and Technology Hellas/Hellenic Institute of Transport, Thessaloniki, Greece, Phone: +30 2310 498 433, E-mail: [jose@certh.gr](mailto:jose@certh.gr)

<sup>4</sup>IMTECH, Netherlands, Phone +31 6 53 207 206, Email: [jaap.vreeswijk@imtech.com](mailto:jaap.vreeswijk@imtech.com)

<sup>5</sup>Centro Tecnológico de la Automoción de Galicia, Phone, +34 986 900 300 Email: [josemanuel.martinez@ctag.com](mailto:josemanuel.martinez@ctag.com)

<sup>6</sup>Volvo Group Trucks Technology (GTT), Phone, +33 426 831 434 Email: [guillaume.vernet@volvo.com](mailto:guillaume.vernet@volvo.com)

### **Abstract**

This paper describes the methodological framework of the COMPASS4D project, which aims to deploy cooperative mobility systems in seven European cities during a one-year full scale pilot. The underpinning approach to data handling, quality controls and analysis is specified, along with the experimental design, describing the split between two separate baseline and operation phases, performance simulation, and quantitative and qualitative data collection. The approach to network performance modelling, which extrapolates the impacts of the deployed COMPASS4D services to whole cities, both at the infrastructure and fleet sides, is outlined. The full paper and presentation will disseminate results from the provision of services under real world conditions, including simulation and network performance modelling activities, showing how Compass 4D services can move beyond trial deployments to extend across whole cities and networks.

**KEYWORDS: cooperative; deployment; evaluation**

### **Introduction**

The COMPASS4D project is funded by the European Commission (EC) under the Competitiveness and Innovation Framework Programme (CIP) to promote the deployment of co-operative mobility systems in real-world conditions in order to improve road safety, increase energy efficiency, and reduce congestion for road transport.

Three services have been implemented and operated in real driving conditions at seven pilot locations: Bordeaux, Copenhagen, Helmond, Newcastle, Thessaloniki, Verona, and Vigo. Data from these deployments is undergoing analysis in order to provide insights into system performance, user acceptance and behaviour, and possible network-wide impacts. The three services are:

ROAD HAZARD WARNING: The Road Hazard Warning (RHW) service aims to reduce incidents by sending warning messages to drivers to raise their attention level and inform them about appropriate behaviour. The advantage of a cooperative RHW service using infrastructure-to-vehicle (I2V) communication is twofold. First, the replacement of expensive vehicle sensors by cooperative technology enables vehicles without such sensors to be aware of road hazards. Second, the service already has an impact with low market penetration as the hazards are detected and announced by the infrastructure.

RED LIGHT VIOLATION WARNING: The Red Light Violation Warning (RLVW) service aims to increase drivers' alertness at signalised intersections to reduce the number of accidents or reduce the impact of accidents in case they still happen. Although the focus of the service is on red light violation it also addresses situations involving emergency vehicles as well as right of way rules at signalised intersections in a more general sense. The advantage of a cooperative RLVW service using I2V communication over conventional repressive solutions is its intervention before, instead of after, an event occurs.

ENERGY EFFICIENT INTERSECTION SERVICE: The Energy Efficient Intersection Service (EEIS) aims to reduce energy use and vehicle emissions at signalised intersections. The major advantage of a cooperative EEIS using I2V communication is the availability of signal phase and timing information (SPaT) in the vehicle. Presenting this information to drivers enables them to anticipate the current and upcoming traffic light state.

Over a 12 month operational phase these services are being piloted by more than 400 different vehicles representing many types of road user (buses, heavy goods vehicles, emergency vehicles, taxis, electric vehicles, private cars) including public (fleet) and private drivers. The relation between the piloted services, the vehicles and the cities is presented in Table 1.

In addition, each city has a suite of road side units deployed in its operational zone to enable V2I communication. More details about the COMPASS4D project can be found in Mitsakis *et al.* 2014.

**Table 1 Overview of COMPASS4D Deployment Characteristics in the Pilot Cities**

Cities	Services					
	RHW		RLVW		EEIS	
Bordeaux	X		X		X	
Copenhagen	X		X		X	
Eindhoven - Helmond	X		X		X	
Newcastle					X	
Thessaloniki	X				X	
Verona	X				X	
Vigo	X		X		X	
Cities	Users					
	Trucks	Private cars	Buses	Emergency vehicles	Taxis	Total
Bordeaux	5	20		6		31
Copenhagen	17	2	87			106
Eindhoven - Helmond	8	6	2	5	10	31
Newcastle		2		11		13
Thessaloniki		7			150	157
Verona		40	10			50
Vigo		8	20	2	10	40

### Evaluation Framework

The evaluation framework is based on the FESTA methodology and draws on experiences of other cooperative systems projects including FREILOT, Drive C2X and COSMO. It defines the tasks and methodologies to be followed for the evaluation of the COMPASS4D services in order to achieve a common understanding and a harmonized approach and thus valid and comparable results.

Six impact areas were defined, which serve to evaluate the three COMPASS4D services.

- Road safety (i.e. accident rates, severity of accidents, severity of injuries);
- Network efficiency (i.e. traffic flows, traffic volumes, travel time, travel time reliability, average vehicle speed, reliability of public transport);
- Environmental impact (i.e. CO<sub>2</sub>, NO<sub>x</sub>, PM<sub>10</sub> emissions, fuel consumption);
- Economic sustainability (i.e. maintenance cost, network monitoring, business models for implementation of cooperative systems);
- Traffic network management (i.e. long-term driver behaviour, travel behaviour, compliance, unintended impacts depending on road type, weather conditions, etc.); and
- Driver specific metrics (i.e. user acceptance, user experience and compliance, workload)

and behavioural change).

The evaluation process is based on the study of the behaviour of the drivers in different use cases with the services active, against their behaviour during the baseline operation (when the services are not switched on). To do that effectively, research questions and hypotheses have been defined, to be studied using performance indicators. These indicators influence which data need to be collected and what measurements have to be recorded.

Each COMPASS4D pilot site has identified site-specific use cases. These were created to understand the local variations of each service, and represent an abstraction level to allow for the evaluation of each use case to be applicable for each pilot site which implemented this use case. These use cases have been described in project deliverables D2.4 *Installation and user guides per site* and in D4.1 *Evaluation framework* and are summarised in the tables below.

**Table 2 Use Cases of RLVW Service**

<b>Event_ID</b>	<b>Use_case</b>	<b>Bordeaux</b>	<b>Copenhagen</b>	<b>Helmond</b>	<b>Newcastle</b>	<b>Thessaloniki</b>	<b>Verona</b>	<b>Vigo</b>
101	Emergency Vehicle Warning	X		X				X
102	Own red light violation warning							
103	Red Light Negation Warning		X	X				X
104	Red Light Violator Warning		X	X				
105	Turning warning – oncoming traffic		X	X				
106	Turning warning – crossing vulnerable road users		X	X				

**Table 3 Use Cases of RHW Service**

<b>Event_ID</b>	<b>Use_case</b>	<b>Bordeaux</b>	<b>Copenhagen</b>	<b>Helmond</b>	<b>Newcastle</b>	<b>Thessaloniki</b>	<b>Verona</b>	<b>Vigo</b>
201	Static use case (event created on back office)							
202	Road Hazard: Stopped vehicle ahead							
203	Road Hazard: Accident/incident ahead	X					X	
204	Road Hazard: Road works (exit) ahead			X				
205	Bridge closed / open							

206	Road Hazard: traffic jam/queue ahead	X		X		X	X	X
207	Road Hazard: road works	X	X				X	
208	Road Hazard: detour ahead		X					
209	Fallen Object							
210	Inverse Direction							
211	Pedestrian							
212	Under speed							

**Table 4 Use Cases of EEIS Service**

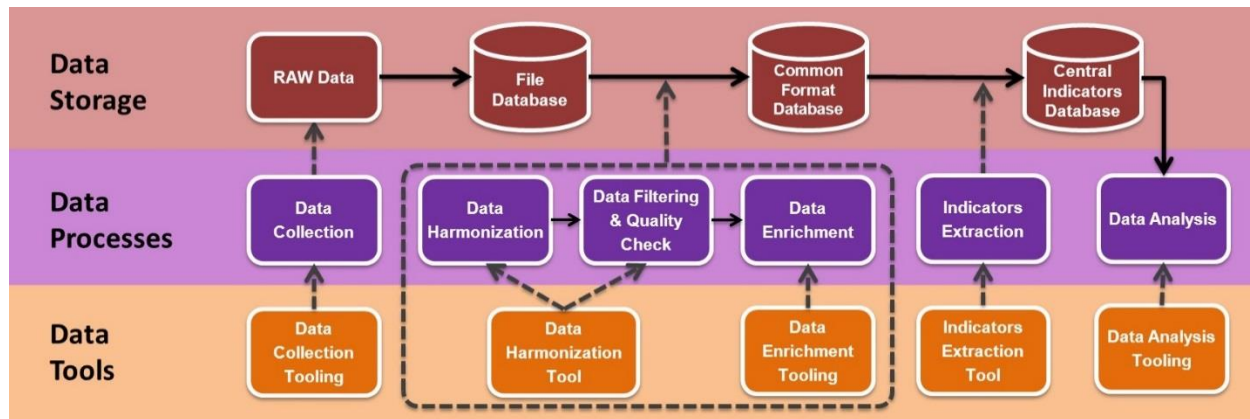
<b>Event_ID</b>	<b>Use_case</b>	<b>Bordeaux</b>	<b>Copenhagen</b>	<b>Helmond</b>	<b>Newcastle</b>	<b>Thessaloniki</b>	<b>Verona</b>	<b>Vigo</b>
301	Green Light Optimal Speed Advisory	X	X	X	X	X	X	X
302	Green priority		X	X	X	X		X
303	Idling stop support	X	X	X			X	
304	Start delay prevention support		X	X				
305	Emergency vehicle priority							
306	Public transport							

### Study Design

The duration of the evaluation phase is 12 months, including 3 months baseline and 9 months full operational with the services switched on. The target users of the Compass 4D project are drivers of buses, emergency vehicles, trucks, taxis, electric vehicles and private cars. The study is therefore targeted at real drivers in real-world driving conditions.

### Data Collection

In order to account for the differences in trial set-up and use of legacy equipment, it was decided that raw data collection would be the responsibility of the pilot sites. This means that the data management procedure has to include three steps that describe the transfer of the data from the pilot sites to the central server. This data chain, shown in Figure 1, includes: data storage, data processes and data tools.



**Figure 1 Data Handling in COMPASS4D**

Raw data to be collected is defined by the performance indicators. Each pilot site implements its own data collection tools making use of different data loggers, and is also responsible for implementing monitoring procedures to ensure the availability and quality of data, based on guidance from COMPASS4D evaluation partners.

Due to the existence of different data loggers a common file format has been defined to facilitate the task of sharing and comparing data between the pilot sites. Sites may collect data in the common format or in alternative formats, later adapting the data into the common format. In both cases a local database is created. This local database can be used for monitoring purposes. Multi-source data may need to be harmonized (and synchronized) in order to avoid errors during the data analysis. The data quality will be checked in terms of: distribution of values, range detection, etc. Ultimately pilot sites can export the data to a file and send this file periodically to the central server. To transfer data from pilot sites' local databases to the central server, a SFTP (Secure File Transfer Protocol) is used. This is a network protocol that provides file access, file transfer and file management functionalities over any reliable data stream. Once the logs are in the central server a tool will process these files to include data in the database. During this process a quality check will also be performed for detecting errors during transmission and other possible errors in the data.

In addition, data will be processed in order to obtain summaries (by trip and by event) and to establish the necessary relationships between tables in the database. This task will facilitate the extraction of indicators and the analysis process.

The final step is to develop a tool to calculate the indicators needed for further analysis and to introduce these indicators into a specific database.

Since each pilot site supplies regular logs in the common data format from its local database to the central server, it was agreed that monitoring tools are developed and hosted on the central server. With this in mind, a web-based interface seems to be the easiest way to provide access to the tool for the monitoring teams of each pilot site.

The functionalities of the monitoring tools are:

- User login (the tool should not be accessible to the whole internet, and the account should also be linked to a list of pilot sites)
- Selection of time period:
  - Pre-sets: week, day
  - Free selection of start / end
- Selection of indicator to monitor
- Export function (e.g. csv, pdf, image)

To avoid pilot site monitoring personnel actively having to log in to the web tool, an automatic email system is provided. This email is sent weekly providing pilot site monitoring information for the previous time period.

With the basic data sets derived from a GPS location and timestamp (date/time, latitude, longitude and altitude) it is possible to calculate the majority of the indicators such as the average speed of a vehicle, the distance travelled per trip, or journey duration.

Before deriving a metric, all the data without a valid timestamp is removed from the datasets before being ordered by timestamp. Data from vehicle loggers, especially if local caching of the data is occurring, can frequently be sent out of order. If the measures for the vehicle need to be created for each individual trip then the data will also need to be separated according to an algorithm that can split up data into individual trips. In previous work this was accomplished using the ignition and flagging a timestamp as being either the start of a trip (if the ignition was turned on) or the end of a trip (if the ignition was turned off).

## **Data Analysis**

### *Analysis structure*

Analysing the data from a single project is normally a relatively simple affair. However, in Compass4D we are analysing data from multiple different pilot sites, with slightly varying data schemas, on different time scales under different use cases. The volume of data, and the speed with which the data is generated, means that we are almost approaching a use case where the phrase “Big Data” would be appropriate.

The philosophy behind the analysis structure is to create a method that will allow for the systematic improvement of all areas of analysis. To achieve this a number of steps were taken:

- 1) The data from each pilot site was unified into a comprehensive data structure at the earliest possible point. If multiple different data schemas are being used (or data sets are being fit into the schema in different ways) then it soon becomes difficult to deal with the multitude of errors that will be thrown up for each individual data type.
- 2) The links between the separate areas of the analysis are designed to be static “anchoring” points which allow for a faster iteration of analysis within each specified analysis area.
- 3) Each section is designed to perform a specific job, ranging from basic file handling and data back-up, detailed generation of performance indicator to front end presentation of the results.

The entire analysis and reporting procedure may be broken up into separate sections, as follows:

### *Data Transfer and File Management*

This is accomplished through a few Python scripts which are designed to automatically transfer the data from the central server to a local file system at the University of Newcastle and store each SQLite file in an appropriate location. Although this step is automated it is also possible to accomplish this using a basic FTP file server (such as FileZilla) if the analysis is to be done in several discrete batches.

### *Performance Indicator Generation*

This is currently both the most time-consuming step in terms of programme run time and in terms of the work required to achieve this step. The data from the SQLite tables typically have multiple different errors across the different sites in addition to systematic variations between the sites with regard to how the data is handled.

The PI generation is currently being handled using the statistical programming environment R. R is a high level language with multiple embedded features that make it possible to evaluate and report on data at a higher pace than with other traditional languages. The development time for creating the PI generation scripts is typically greatly reduced when compared to other languages.

### *Visualisation and Interactive Analysis*

The third major section is the creation of a series of tools that will not only allow the PIs to be visualised, but will also allow for a suite of other functions such as statistical significance testing, selective statistical reporting, data download and comparative visualisations between the different phases for each pilot site.

It is thought that by creating a user friendly tool to allow for direct visualisation and statistical tests on the data, it will enable sites to create the detailed questions that they need to answer the viability of their own use cases. In particular it is thought that operating in this way will allow greater access to the data, and at an earlier timeframe, for those tasked with assessing the cost benefit analysis and other business case-related elements of the project; in other words to provide a full impact assessment of COMPASS4D.

The tool for the visualisation and analysis has been created using R and Shiny:

- R (as has been previously mentioned) is a programming environment that is ideally suited for large scale analysis of data sets including subsequent visualisation and more detailed modelling using that data.
- Shiny is a web application framework that wraps around R to generate interactive web pages linked to an underlying series of statistical scripts. It allows for R scripts, and the visualisation generated from them, to be dynamically altered to enable users to explore the data generated in the Compass4D project. The tool may either be accessed through a web browser or run locally on a local installation of R.

The final version of this tool contains the following features (amongst others):

- 1) Significance testing including variable comparison. To allow for the effects of exogenous variables causing any observed changes in the driving patterns of the COMPASS4D users.
- 2) Spatial mapping of features of interest. To quickly spatially identify the areas of the COMPASS4D zones that exhibit interesting features.
- 3) Automated report generation based on criteria. This will automatically generate 1-2 page reports on selected variables under selected conditions. For example, a report detailing the statistics of a particular junction during weekends could be requested.
- 4) Full and partial data downloads, with aggregations. This will download the data from the PI database to a csv file using a variety of typical database style operations (joins etc.) to deliver the data in the style needed.

### **Early Results**

The following section shows a few short examples of analysis on the existing data from the Pilot Sites. It should be stressed that this data is from an early iteration of the PI generation and as such is not at the same standards of data quality as will be evidenced in the final analysis procedures and reporting.

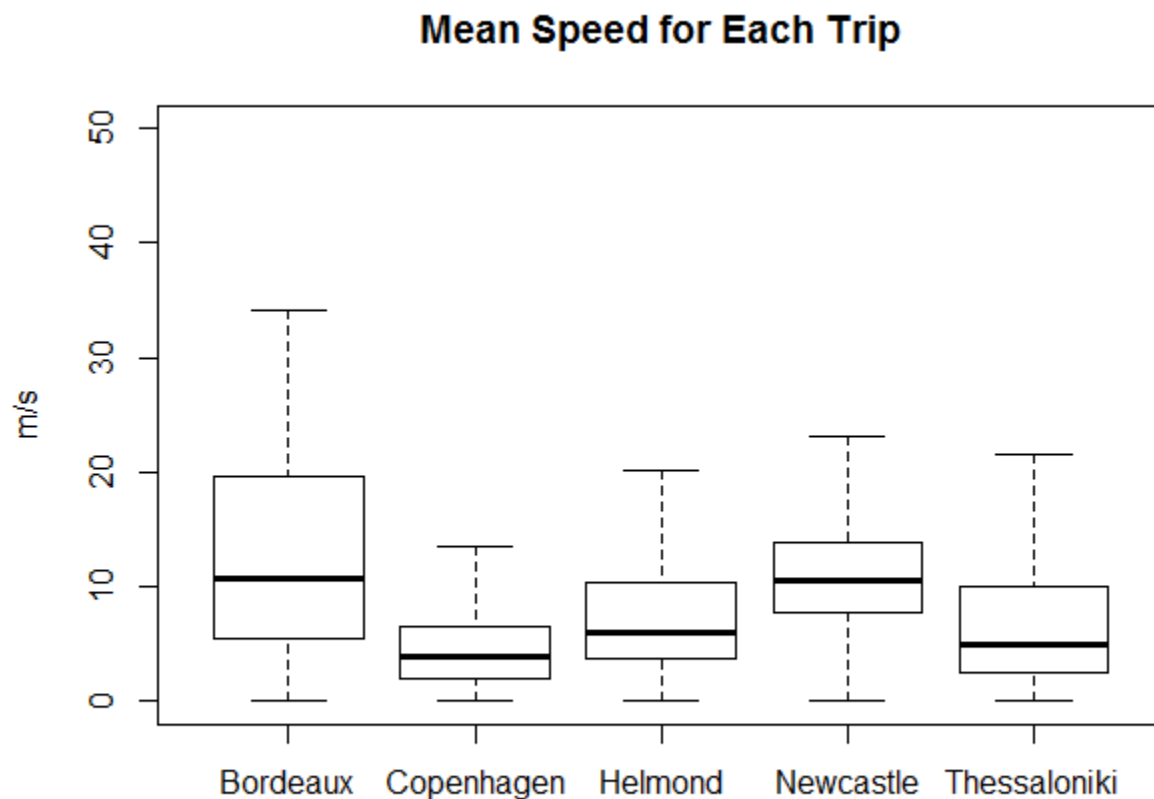


**Table 5 Summary of Statistics by 31<sup>st</sup> May 2015 (5 sites analysed)**

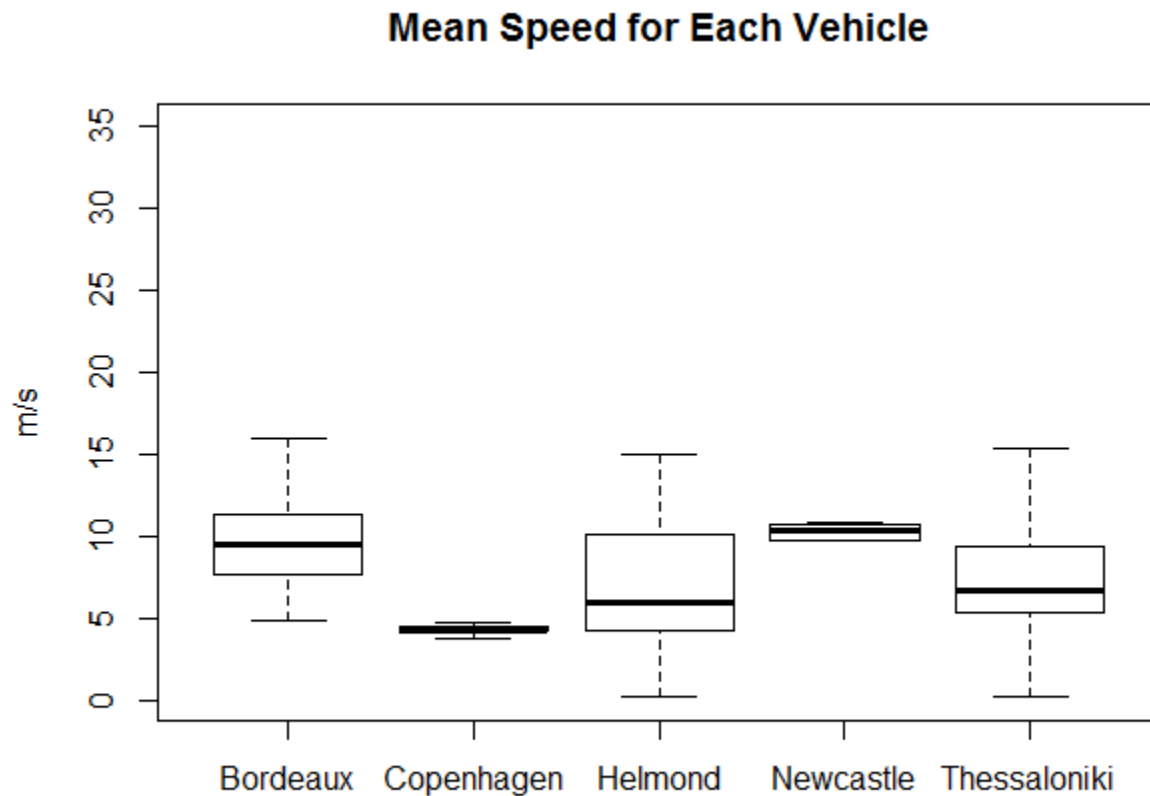
	Mean Distance	Mean Speed	SD Speed	SD Duration
Bordeaux	673	8.64	7.25	161
Copenhagen	361	6.06	4.78	140
Helmond	678	8.18	7.20	149
Newcastle	644	8.92	7.40	43
Thessaloniki	577	5.47	5.00	154

Table 5 shows a few brief examples of the type of statistics that are useful to publish. In general it would not be valid to compare data between the different pilot sites and instead each statistic would be aggregated against either each vehicle ID or each intersection ID. Some statistics do stand out in this table, especially the low standard deviation of duration for Newcastle.

In Figure 2 we can see how the mean speed for each recorded trip varies between the five pilot sites for which data had been analysed. In this graph the box represents  $\pm 1$  S.D. from the mean (thick black line) with the whiskers representing  $\pm 2$  S.D.

**Figure 2 Box Plot Showing the Means and Ranges of the Speeds Measured in the Pilot Sites**

There is a wide variation in the standard deviation points amongst the five sites, and in general a wide range of maximum speeds observed within the sites. This is entirely expected as the sites exhibit a wide range of different characteristics including speed limits and road capacity. However, if the sites are to be compared with each other then it is necessary to gain a good understanding of exactly how the data within each site varies. Each site will have a difference in number of vehicles as well as different vehicle fleet composition. An increase in average speed of 1m/s will not be significant for a car, but for a slow moving bus it could lead to a large reduction in emissions.



**Figure 3 Box Plot Showing the Means and Ranges of the Speeds Measured in the Pilot Sites with Data Aggregated by Vehicle ID**

In Figure 3 it can be seen that although the sites all exhibited a reasonably similar standard deviation in their whole population statistics, there is a marked reduction in the standard deviation of the vehicle statistics in both Copenhagen and Newcastle. There are many possible explanations for such a discrepancy, but one explanation could be in the driving patterns of the drivers. If the drivers in Newcastle and Copenhagen are driving uniformly through all the intersections, then there is no great reason to expect that their average speed would vary. But if they were only driving through certain intersections, then the speed variation imposed by those intersections would lead to a splitting of the population. In effect, each driver in Newcastle/Copenhagen is acting as uniform sample of the intersections, whilst those in Bordeaux, Helmond and Thessaloniki are acting as a

selective sample. This could be checked by examining the variation of intersections visited by each vehicle across each site, and then forming a model which seeks to understand the level of variance that can be explained by the number of intersections visited.

Completing this sort of detailed analysis will be the final task of the data analysis. It will:

- 1) Examine the data to test for a difference between the phases of the trial
- 2) Check for statistical significance
- 3) Construct models to attempt to look for explanations of the differences, beyond simply phase changes

If, after moving through each of these steps, we have a statistical difference between the data that cannot be explained then we may state firmly that COMPASS4D has improved the driving experience in the pilot sites.

### **Measuring Network-Wide Impact of Cooperative ITS Services**

Although the number of participants and cooperative systems at each pilot site is limited, it is believed that the use of the proposed evaluation methodologies will provide valid results. A key component of the evaluation approach is evaluation of network performance, which extrapolates the impacts of the deployed COMPASS4D services to quantify the benefits to whole networks, cities, and participating organisations. The network under study can comprise city streets and/or interurban roads, depending on the characteristics of the service to be examined. This too seeks to inform the business model. The provision of network-wide data for different network geometries and demand densities (expressed in number of trips per surface/area unit) will be useful for evaluating network performance for any European city, with different network geometries and demand levels. This approach will significantly enrich the end-result of the project, since customised estimations of benefits and costs could be provided for any city.

Three methodologies will be implemented:

- A techno-economic assessment based on data envelopment analysis and using data from large scaling network simulation (EEIS);
- An analytical methodology based on continuous approximations and geometrical probabilities (EEIS);
- An empirical statistical methodology based on desktop research of the impacts of each service under various traffic conditions and network types (RHW/ RLVW).

### **Conclusion**

The pilot sites are currently engaged in their full operational phase, with data collection for evaluation purposes now scheduled to run until the end of November 2015. This paper presents early indicative results from data analysis up to the end of May 2015, along with a more detailed description of the analysis procedure. The presentation will present more complete and up-to-date results from all evaluation phase based on data analysis up to the end of August. Results of simulation and modelling activities showing how COMPASS 4D services can be extrapolated across network levels will also be more fully described. Overall, it is envisaged that the results to be presented will prove the likely safety and energy efficiency benefits of COMPASS 4D services to all relevant stakeholders.

## References

1. AMITRAN project - <http://www.amitran.eu/>
2. CIP - <http://ec.europa.eu/cip/>
3. COSMO project - <http://www.cosmo-project.eu/>
4. CVIS project - <http://www.cvisproject.org/>
5. DRIVEC2X project - <http://www.drive-c2x.eu/project>
6. FREILOT project - <http://www.freilot.eu/>
7. Last revision of FESTA methodology available at [http://www.fot-net.eu/download/festa\\_handbook\\_rev4.pdf](http://www.fot-net.eu/download/festa_handbook_rev4.pdf)
8. Mitsakis E., Salanova Grau J.M., Aifandopoulou G., Tona P., Vreeswijk J. Somma G., Blanco R., Alcaraz G., Jeftic Z., Toni A. and Blom G. 2014, *Large scale deployment of cooperative mobility systems in Europe: COMPASS4D*, The third International Conference on Connected Vehicles and Expo, ICCVE 2014, Vienna, Austria
9. SmartWay project - <http://www.smart-way.mobi/>